Predicting Water Pump Functionality in Tanzania

BY LEWIS NGUNJIRI

Overview and Objectives

Overview:

 - Predicting water pump functionality in Tanzania

 - Importance of maintaining functional water pumps

Objectives:

- Analyze factors affecting water pump performance

- Develop a predictive model for water pump functionality

Background:

Water access is a critical issue in Tanzania, affecting millions of people. Despite significant investments in water infrastructure, many communities still face challenges in accessing clean and reliable water. The functionality of water pumps plays a crucial role in ensuring that communities have continuous access to this vital resource. However, a significant number of water pumps become non-functional over time due to various reasons such as poor maintenance, environmental conditions, and lack of resources.

Context:

- Over 30% of water pumps in Tanzania are non-functional

- Access to clean water is critical for health and development

Challenges:

- Limited resources for maintenance and repairs

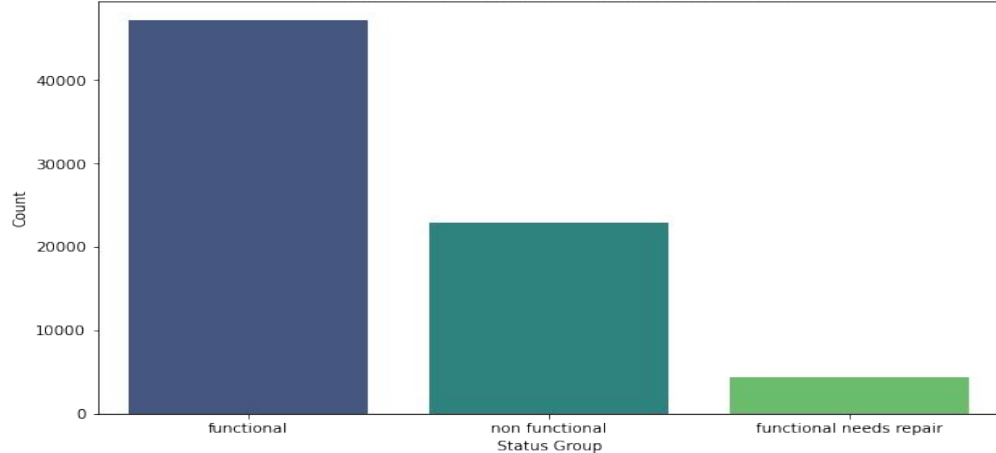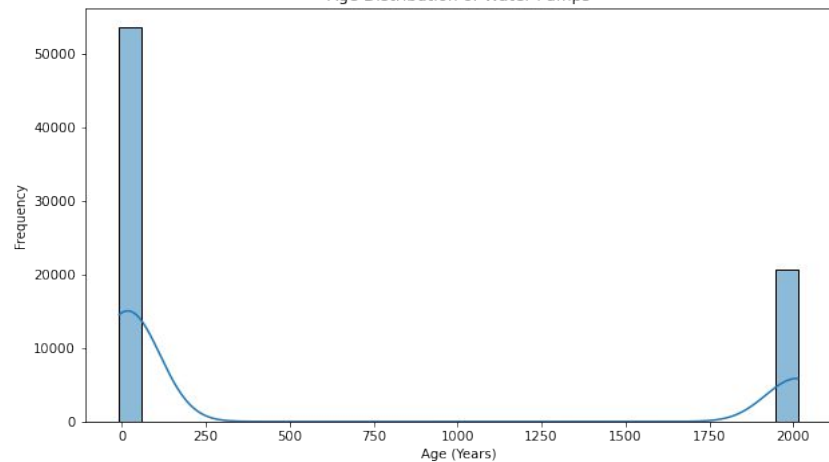- Variability in pump installation and quality

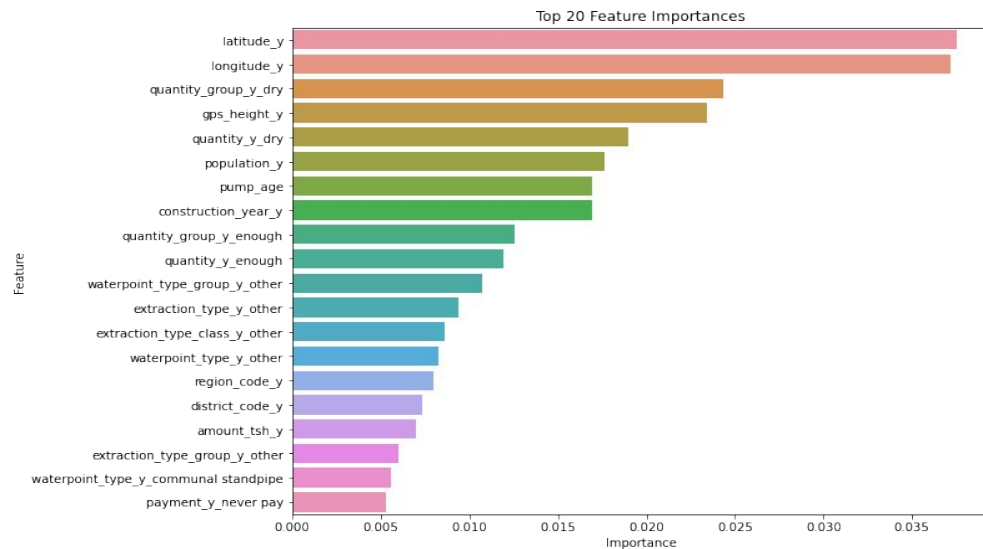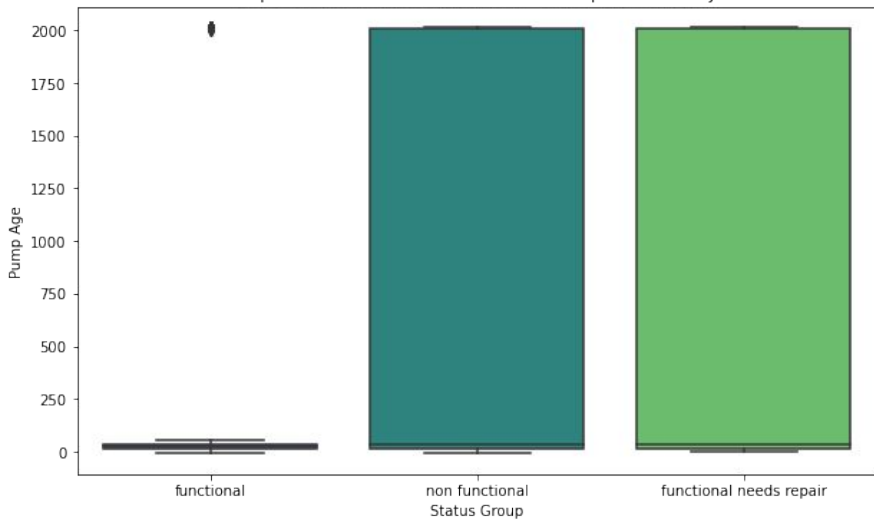# Exploratory Data Analysis (EDA)

Key Findings:

- Distribution of pump statuses

- Correlation between features

- Geographical distribution of pump functionality

Age Distribution of Water Pumps

Impact of Installation Year on Water Pump Functionality

Top 20 Feature Importances

Explanation of Visualizations

Distribution of Water Pump Functionality Status:

This bar plot shows the distribution of the different functionality statuses (functional, needs repair, non-functional) of water pumps. It helps in understanding the overall health of water pumps in the dataset.

Age Distribution of Water Pumps:

This histogram shows the distribution of the ages of water pumps. It helps in understanding how the age of pumps varies and how many pumps are relatively new or old.

Impact of Installation Year on Functionality:

This box plot visualizes the relationship between the installation year (age) of water pumps and their functionality status. It helps in identifying trends, such as whether older pumps are more likely to be non-functional.

Model Building

Models Tried:

- Random Forest, Gradient Boosting,

Logistic Regression Training Process:

- Split data into training and test sets

- Hyperparameter tuning using GridSearchCV

Results and Discussion

Functional Class:

Precision (0.87): Out of all instances predicted as functional, 87% were correctly classified. Recall (0.93): Out of all actual functional pumps, 93% were correctly classified by the model. F1-Score (0.90): The harmonic mean of precision and recall for the functional class is 90%.

Functional Needs Repair Class:

Precision (0.60): 60% of instances predicted as needing repair were correctly classified. Recall (0.35): Only 35% of actual pumps needing repair were correctly classified by the model. F1-Score (0.44): The F1-Score, which balances precision and recall, for this class is 44%.

Non-Functional Class:

Precision (0.85): 85% of instances predicted as non-functional were correctly classified. Recall (0.79): 79% of actual non-functional pumps were correctly classified by the model. F1-Score (0.82): The F1-Score for the non-functional class is 82%.

Factors to Consider for Best Results on Pump Performance and Reliability

1.Pump Type and Specifications: Choose pump types that are suited to the specific needs and conditions of the area (e.g., hand pumps for shallow wells, submersible pumps for deep wells).

2.Water Quality and Source Protection: Ensure that the water source is protected from contamination and that the pump is equipped to handle the specific water quality (e.g., pumps resistant to corrosion for saline water).

3.Infrastructure and Accessibility: Develop infrastructure around the pump site, such as access roads and storage facilities, to facilitate easy maintenance and repair work.

4.Environmental Impact Assessment: Conduct thorough environmental impact assessments before installation to ensure that the pump does not adversely affect local ecosystems.

Recommedations and Conclusion

To improve the performance and reliability of water pumps across Tanzania, consider the following:

1.Site Selection and Installation Geographical and Environmental Factors:

Elevation and Terrain: Install pumps in areas where the geographical features do not pose a risk of flooding or excessive wear due to rough terrain. Soil Quality: Ensure the soil around the installation site is stable and not prone to erosion or landslides. Water Table Levels: Select sites with a stable and sufficient water table to ensure continuous water availability.

2.Regular Maintenance and Monitoring

Scheduled Maintenance: Implement a routine maintenance schedule to check and service pumps regularly, preventing minor issues from becoming major problems. Remote Monitoring Systems: Utilize remote monitoring technology to track pump performance in real-time, enabling quick response to any emerging issues.

3.Data-Driven Decision Making

Data Collection: Continuously collect data on pump performance, environmental conditions, and usage patterns. Predictive Analytics: Use machine learning models to predict potential failures and plan proactive maintenance.

4.Quality of Materials and Installation Practices

High-Quality Materials: Use durable materials for pump components to reduce wear and tear and increase longevity. Skilled Labor: Ensure that the installation is carried out by skilled technicians to prevent early failures due to poor workmanship.

Recommendations Summary

Practical Steps:

- Prioritize maintenance for older pumps

- Standardize installation practices

- Train and monitor installers

Actions:

- Allocate resources for regular maintenance

- Develop a centralized database to track pump performance and maintenance